Making NOAA´s data more accessible to the scientific community

Chris Fox
Director, National Geophysical Data Center

CLASS Users' Workshop
August 11, 2005

History of CLASS@NGDC

- NRC program review recommended OSDPD return to its basic mission of data processing and product generation… not data archive
- In March 2004, RGR approached CGF to relocate Suitland CLASS site to NGDC
- Move was approved by NESDIS on 2/8/05
- Efforts are underway to install CLASS hardware and personnel at NGDC by Spring, 2006

Why NGDC?

- NGDC's Mission is Data Archive, Access, and Assessment
- Excellent Infrastructure in Boulder
  - New facility, connectivity (Lambda Rail, BRAN)
- Connection with the Boulder-area Scientific Community
  - UCAR/NCAR/UNIDATA, CU, NOAA Labs, Commercial
- Local expertise in Metadata, Access Tools, Data Management, Systems Administration

What CLASS Brings to Boulder

- All NOAA satellite data (and potentially nearly all NOAA data) will be available to the Boulder scientific community at LAN bandwidths via the Boulder Research and Administration Network (BRAN)
- These same data will available to the non-Boulder community at Lambda-rail bandwidth from two, redundant sites

Fox's Vision for NOAA Data Management

All NOAA environmental data will…

- reside in secure archives conforming to NARA and COOP standards
- be maintained at the highest standards of scientific data stewardship
- be searchable using advanced data discovery tools, to facilitate interdisciplinary studies
- be accessible through a common portal available to the scientific community, commercial sector, and general public based on advanced access tools
- be made accessible to broader data integration efforts, such as IOOS and GEOSS

Guiding Principles

All NOAA environmental data will…

- Reside in secure archives conforming to NARA and COOP standards
  - CLASS will provide a dual-site secure archive for all NOAA data. Large-array data sets (such as NPOESS, GOES-R, NEXRAD) will be directly archived; Ultimately, CLASS would provide archival services for other data sets from NOAA National Data Centers and NOAA Centers of Data
- Maintain the highest standards of scientific data stewardship
  - NOAA data must be managed by the organizational components that collect and use the information
  - Tools must be provided to all Data Managers for effectively using CLASS while not compromising CLASS's integrity
  - Data quality and lineage tracking metadata must be provided

Guiding Principles

All NOAA environmental data will…

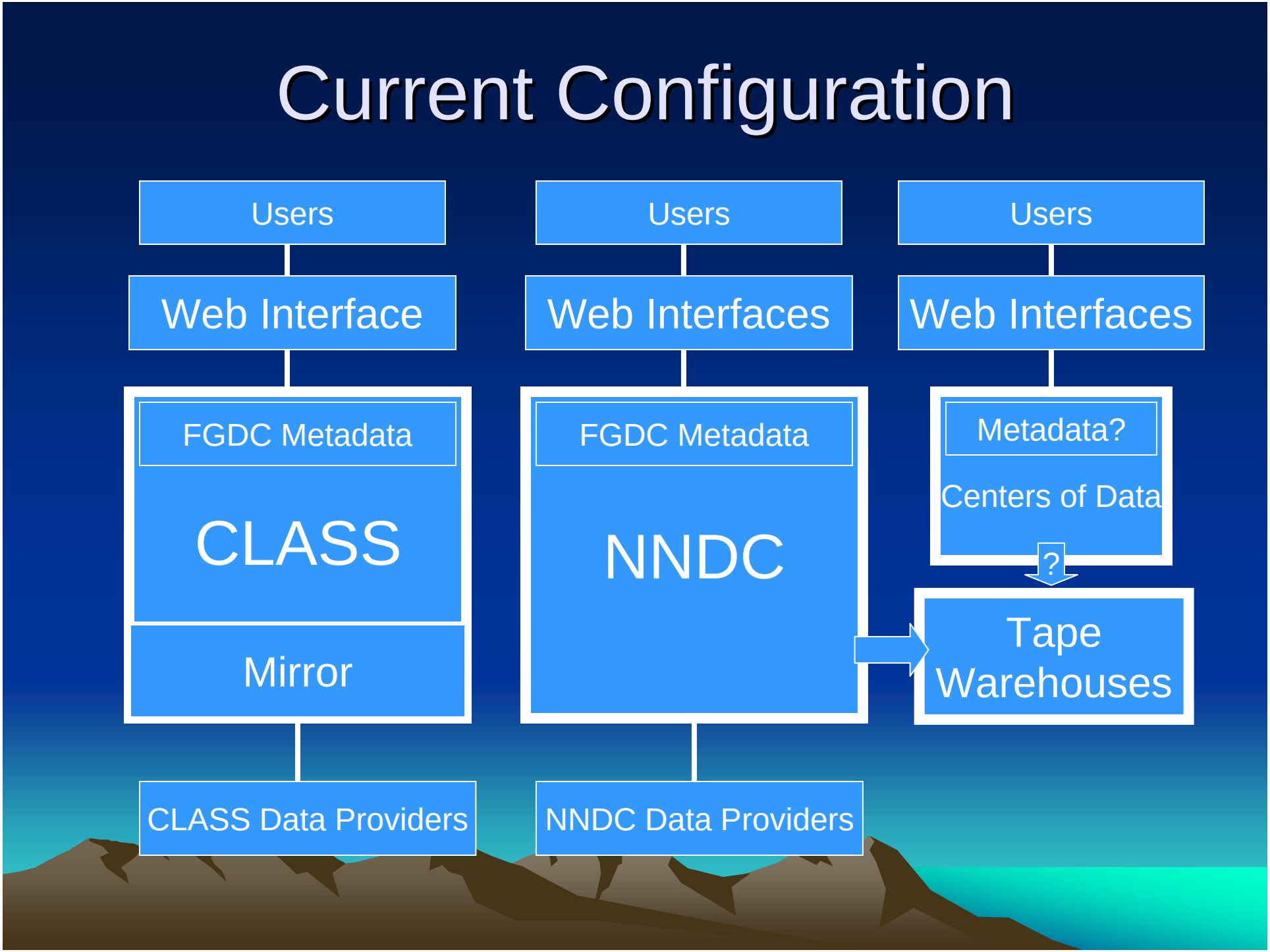
- Be searchable using advanced data discovery tools, to allow interdisciplinary studies
  - The NOAA Metadata Manager and Repository (NMMR) will provide a common “library catalog” for all NOAA holdings; contributors to CLASS will be obligated to maintain their metadata catalog within the NMMR
- Be accessible through a common portal available to the outside community
  - “NOAA Server” will evolve to allow access to data discovery tools, underlying data bases, and a suite of access tools for data analysis and display, including Application Program Interfaces (APIs) and web services to allow machine-machine access and encourage development of value-added products by commercial, academic, and other non-NOAA organizations

Guiding Principles

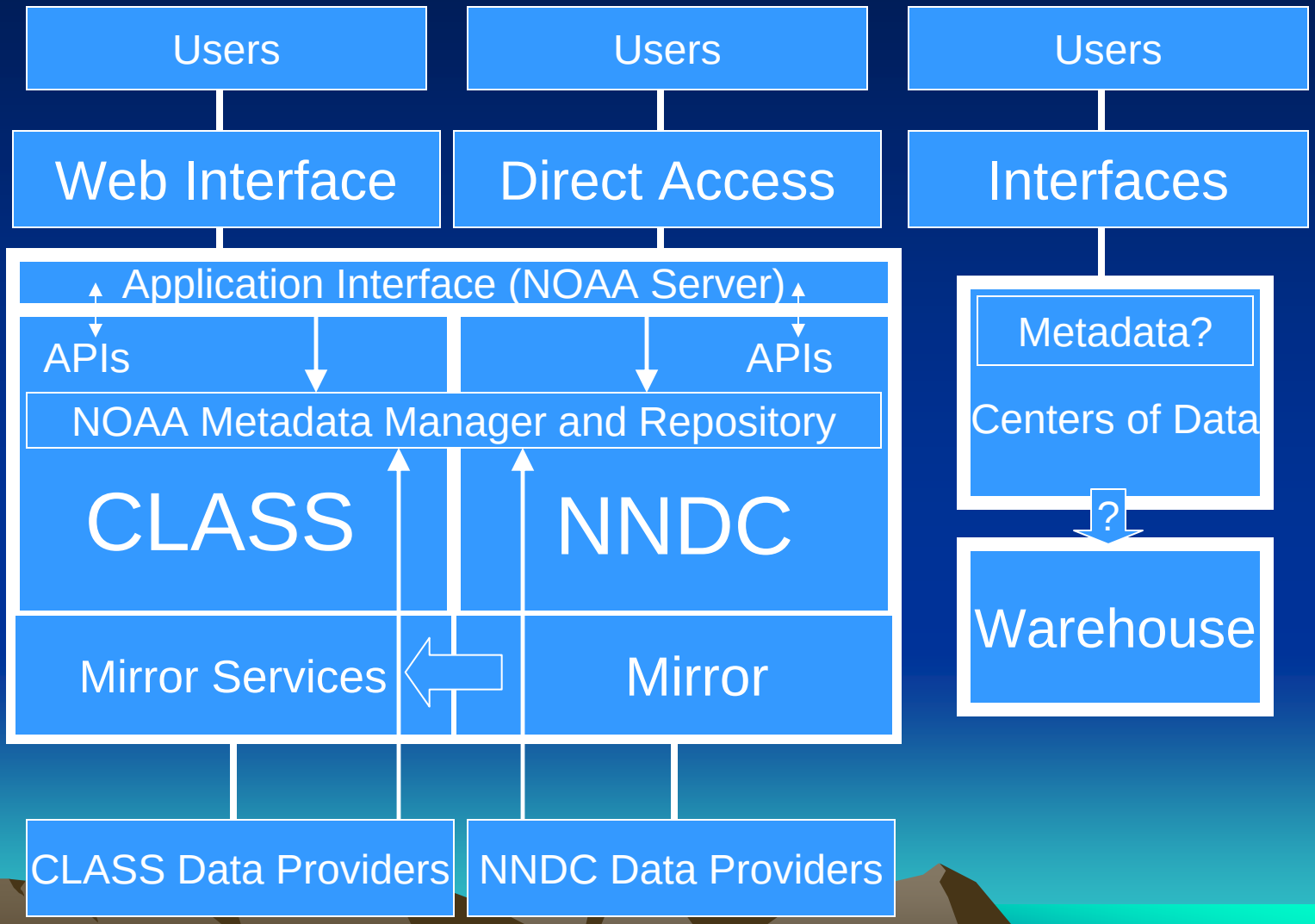
All NOAA environmental data will…

- Be made accessible to broader data integration efforts, such as IOOS and GEOSS
  - NOAA data management systems must be built using accepted international data, metadata, and data transport standards
  - As NOAA web services become available, they will be registered on global data services catalogs to become integrated with GEOSS

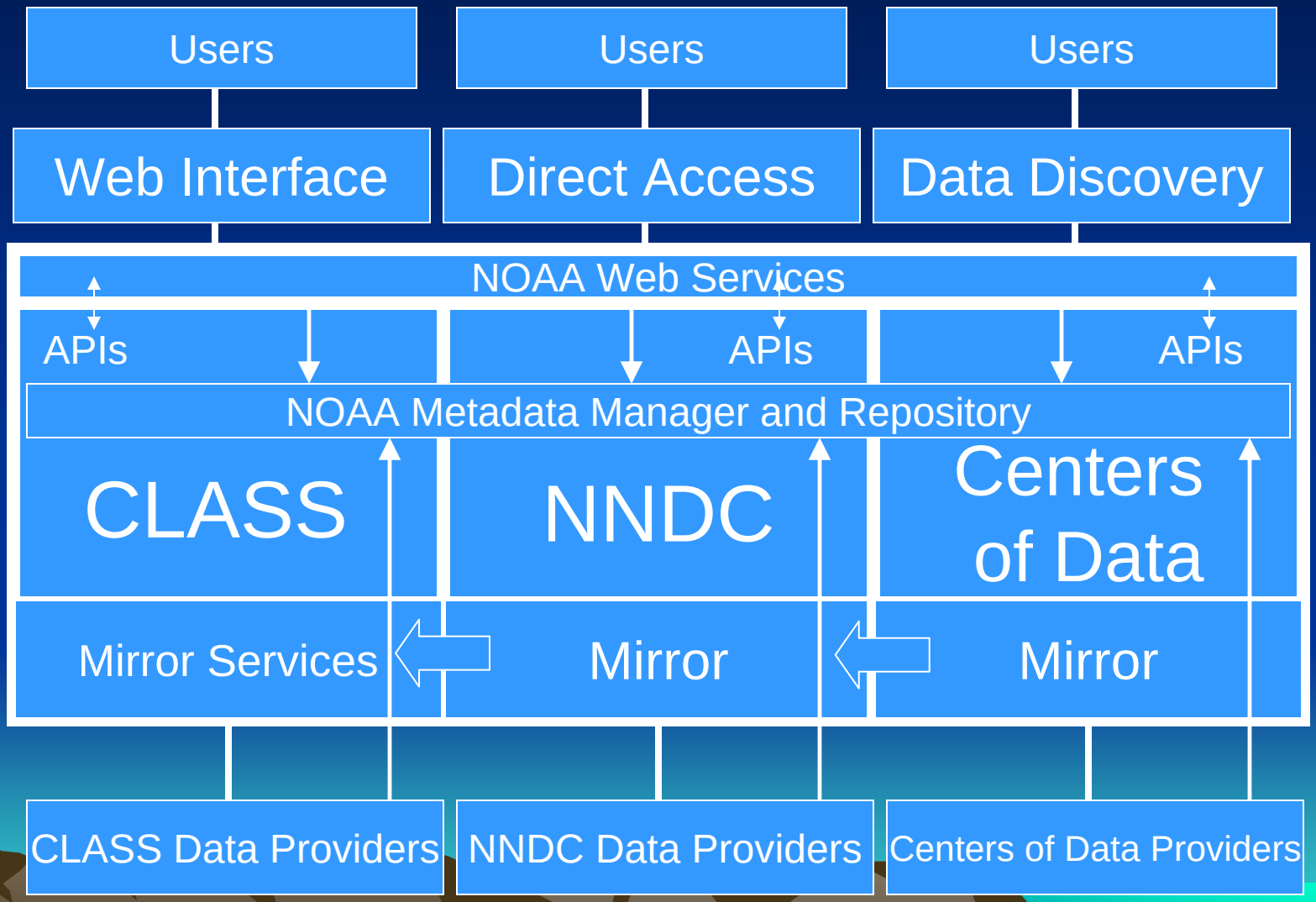
Current Configuration



Stage-2 Configuration



Final Configuration



Summary

- CLASS offers broad-band access to NOAA's large-array data sets from a secure dual-site environment.
- A NOAA-wide data management system (DMS) can evolve from CLASS by integrating with other NOAA data archives and ultimately with NOAA research and other data sets not currently archived
- CLASS will provide a full DMS for large-array data sets, but can also provide secure archival services and access to other NOAA data providers who participate in the NMMR and NOAA-Server
- This semi-distributed architecture, with central data and metadata archives built on international standards, is consistent with the IOOS/DMC architecture and other efforts and will allow future integration of NOAA systems into GEOSS